

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 8, 2002, 21:05:24 ; Search time 8498.8 Seconds  
(Without alignments)  
30.345 Million cell updates/sec

Title: US-09-851-670-2

Perfect score: 24  
Sequence: 1 cgacaacatggaacacagctgcgc 24

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 111874

Minimum DB seq length: 0  
Maximum DB seq length: 60

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em.estfun:\*  
2: em.esthum:\*  
3: em.estin:\*  
4: em.estom:\*  
5: em.estpl:\*  
6: em.estba:\*  
7: em.estro:\*  
8: em.estov:\*  
9: em.htc:\*  
10: gb.est1:\*  
11: gb.htc2:\*  
12: gb.htc:\*  
13: gb.gss:\*  
14: em.gss.fun:\*  
15: em.gss.hum:\*  
16: em.gss.inv:\*  
17: em.gss.pln:\*  
18: em.gss.pro:\*  
19: em.gss.prod:\*  
20: em.gss.vtl:\*  
21: em.gss.other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14.6	60.8	40	10	A1619480 ty38g04.x
2	14.4	60.0	50	10	AU104403 AU104403
3	14.4	60.0	55	10	A1357794 qu98e04.x
4	14	58.3	45	13	A2662129 IM0541J06
5	13.8	57.5	49	13	A2810540 2M0076E08
6	13.6	56.7	29	13	A2831178 2M0110I18
7	13.4	55.8	31	10	A1746720 u106f05.y
8	13.4	55.8	46	10	A1026096 ov94n09.s
9	13.4	55.8	53	13	A2576214 AST-TD13S
10	13.4	55.8	55	11	R60882 yH08f11.r1
11	13.2	55.0	27	13	A2795938 2M0051K23
12	13.2	55.0	38	11	C21091 HUMGS000260

13	13.2	55.0	59	10	A666844 vm5g11.s
14	13.2	55.0	60	10	A1996547 701667389
15	13	54.2	34	10	AA476000 vb26a02.r
16	13	54.2	50	10	AU104180 AU104180
17	13	54.2	50	10	AU104183 AU104183
18	13	54.2	50	10	AU104194 AU104194
19	13	54.2	54	11	T50983 yb71c06.s1
20	13	54.2	54	13	A2783216 2M0024011
21	12.8	53.3	50	11	BG939021 cm30a04.y
22	12.8	53.3	52	13	TA309602P
23	12.8	53.3	53	10	AA528415
24	12.8	53.3	54	5	AW284534
25	12.8	53.3	58	11	W41141 mc38h02.r1
26	12.6	52.5	60	10	AV518812
27	12.6	52.5	25	10	AV518812
28	12.6	52.5	40	10	A1620546
29	12.6	52.5	44	13	A1972180 w763e10.x
30	12.6	52.5	44	13	AZ798809 2M0055L16
31	12.6	52.5	52	10	A1000127 os43a10.s
32	12.4	51.7	38	10	BG273140 na134f09.
33	12.4	51.7	46	10	AV836261
34	12.4	51.7	48	10	A1285380
35	12.4	51.7	49	10	HS0003904
36	12.4	51.7	52	10	A1084829 ow89c12.s
37	12.4	51.7	52	10	A1284802 qu11h12.x
38	12.4	51.7	52	10	A1461385 fP44g03.x
39	12.4	51.7	52	13	AZ852502 2M0155D23
40	12.4	51.7	55	10	AA811401
41	12.4	51.7	57	10	AV543344
42	12.4	51.7	57	13	AV543344
43	12.4	51.7	58	10	AW247684 2820293.5
44	12.4	50.8	60	10	A1024966 cv39b10.x
45	12.2	50.8	21	13	AZ463806 1M0272G19
			23	13	AZ336220 1M0066G19

#### ALIGNMENTS

RESULT 1  
A1619480  
LOCUS ty38g04.x1 NCI CGAP uc2 Homo sapiens cDNA clone IMAGE:2281398 3'  
DEFINITION similar to WP:W03D2.1 CE14506 ; contains element MER22 repetitive element ; , mRNA sequence.

ACCESSION A1619480.1 GI:4628606  
VERSION A1619480  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps-rt@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:  
www-bio.llnl.gov/bdrrp/image/image.html

Trace considered overall poor quality  
Insert Length: 2989 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 1  
POLYA-No.

FEATURES Location/Qualifiers

source	1. 40	/organism="Homo sapiens"
	/db_xref="taxon:9606"	
	/clone="IMAGE:2281398"	
	/clone_lib="NCI_CGAP_Ut2"	
	/tissue_type="moderately-differentiated endometrial adenocarcinoma, 3 pooled tumors"	
	/lab_host="DH10B"	
	/note="Organ: uterus; Vector: pCMV-Sport6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.85 Kb. Life Technologies catalog #: 11539-012"	
BASE COUNT	12 a 20 c 8 g 0 t	
ORIGIN		
Query Match	60.8%; Score 14.6; DB 10; Length 40;	
Best Local Similarity	81.0%; Pred. No. 6.4e+04;	
Matches	17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	4 caaatggaataaacagctgcgc 24	
Db	1 CAAAGGCAAAAACCGCGCC 21	
RESULT 2		
LOCUS	AU104403 50 bp mRNA EST 05-APR-2001	
DEFINITION	AU104403 Sgano Homo sapiens cDNA library Homo sapiens cDNA clone	
ACCESSION	HEP11308, mRNA sequence.	
VERSION	AU104403	
KEYWORDS	AU104403.1 GI:13553924	
SOURCE	EST.	
ORGANISM	human.	
	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 50)	
AUTHORS	Suzuki,Y., Tsunoda,T., Yaita,H., Mizushima-Sugano,J., Sese,J., Hata	
	,R., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo	
	,K., Suyama,A. and Sugano,S.	
TITLE	Fine structural analysis of transcription start sites of human	
JOURNAL	mRNA using full-length enriched and 5'-end enriched cDNA libraries	
COMMENT	Unpublished (2001)	
	Contact: Yutaka Suzuki	
	Department of Virology	
	Institute of Medical Science, University of Tokyo	
	4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan	
	Email: yusuzuki@ims.u-tokyo.ac.jp	
	Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano	
	S., Construction and characterization of a full length-enriched and	
	a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).	
FEATURES	Location/Qualifiers	
source	1..50	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="HEP11308"	
BASE COUNT	5 a 12 c 19 g 14 t	
ORIGIN		
Query Match	60.0%; Score 14.4; DB 10; Length 50;	
Best Local Similarity	75.0%; Pred. No. 7.4e+04;	
Matches	18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
QY	1 cgacaaatggaataaacagctgcgc 24	
Db	24 CGACAAATCGCCCAACAGCATGCC 1	
RESULT 3		
	TI357794	

LOCUS	AI357794	55 bp	mRNA	EST	06-JAN-1999
DEFINITION	gu98d04.x1 NCI_CGAP Gas4 Homo sapiens cDNA clone IMAGE:1980126 3' similar to TR.Q16038 Q16038 PBM1 PROTEIN ;contains TARI.t2 TARI repetitive element ; mRNA sequence.				
ACCESSION	AI357794				
VERSION	AI357794.1	GI:4109415			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 55)				
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.				
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://www.bio.lnl.gov/bbrp/image/image.html">www.bio.lnl.gov/bbrp/image/image.html</a>				
FEATURES	source				
	Trace considered overall poor quality Seq primer: -40UP from Glbco High quality sequence stop: 1. location/Qualifiers 1..55 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:1980126" /clone_1id="NCI_CGAP_Gas4" /tissue_type="poorly differentiated adenocarcinoma with slight ring cell features" /lab_host="DH10B" /note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.69 kb. Life Technologies catalog #: 11549-011"				
BASE COUNT	14 a 23 c 18 g 0 t				
ORIGIN					
	Query Match 60.0%; Score 14.4; DB 10; Length 55; Best Local Similarity 75.0%; Pred. No. 7.3e+04; Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;				
OY	1 cgacaaatggaanaacagctgcc 24                          Db 18 CGCCAAAGGAACCCGCCCCC 41				
RESULT 4					
AZ662129/c					
LOCUS	AZ662129 45 bp DNA GSS 14-DEC-2000				
DEFINITION	IIM054JJOEF Mouse 10kb plasmid UNGC1M library Mus musculus genomic clone UNGC1M054JJO6 F, DNA sequence.				
ACCESSION	AZ662129				
VERSION	AZ662129.1	GI:11799275			
KEYWORDS	GSS.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 45)				
AUTHORS	Dunn,D., Aoyagi,A., Barber,M., Beacom,T., Duvall,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.				

REFERENCE  
AUTHORS

1 (bases 1 to 49)  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,  
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.

REFERENCE  
AUTHORS

Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Murinae: Mus  
1 (bases 1 to 29)  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly

**TITLE** 'M', Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A. and Wright, D., Weiss, R.  
**JOURNAL** Mouse whole genome scaffolding with paired end reads from 10kb  
**COMMENT** Plasmid Inserts  
 Unpublished (2000)  
 Contact: Robert B. Weiss  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunne@genetics.utah.edu  
 Insert Length: 10000  
 Plate: 0110  
 Seq primer: CACACGAGAAACAGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 29.

# FEATURES

**SOURCE** 1. 29  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /strain="C57BL/6j"  
 /db\_xref="taxon:10090"  
 /clone="U082M0110118"  
 /clone\_lib="Mouse 10kb plasmid U082M0110118"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6j (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (911473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

# BASE COUNT

8 a 7 c 6 g 8 t

**Query Match** 56.7%; Score 13.6; DB 13; Length 29;  
**Best Local Similarity** 80.0%; Pred. No. 1.7e+05;  
**Matches** 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 5 aaatgaataaacagctgcg 24  
 Db 2 AAATGTGATTAACAGCTGCC 21

**RESULT** 7  
**LOCUS** A1746720 31 bp mRNA EST 22-JUN-1999  
**DEFINITION** u106105.y1 Sugano mouse kidney m1a Mus musculus cDNA clone IMAGE:2065281 5' similar to TR:013953 O13953 HYPOPHYSICAL 39.1 KD PROTEIN C23H4.12 IN CHROMOSOME 1.; mRNA sequence.  
**ACCESSION** A1746720  
**VERSION** A1746720.1 GI:5124984  
**KEYWORDS** EST.  
**SOURCE** house mouse.  
**ORGANISM** Mus musculus  
**REFERENCE** 1 (bases 1 to 31)  
 Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae; Mus.

**AUTHORS** Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.  
**JOURNAL** The WashU-NCI Mouse EST Project 1999  
**COMMENT** Unpublished (1999)  
 Contact: Marra M/WashU-NCI Mouse EST Project 1999  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@wustl.edu  
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:994469

# FEATURES

**SOURCE** 1. 31  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /strain="C57BL/6j"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:2065281"  
 /clone\_lib="Sugano mouse kidney m1a"  
 /sex="female"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Organ: kidney; Vector: pME18S-FL3; Site: 1: DraIII (CAGCTGCTG); Site: 2: DraIII (CAGCTGCTG); 1st strand cDNA was primed with an oligo(dT) primer (ATGCGCCCTTTTCTTTTCTTTT); double-stranded cDNA was ligated to a DraIII adaptor (TGTTCGCTTACTGG), digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CAGCTGCTG, 3' site CAGCTGCTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCGCTTAAGCTGCG and 3' end primer CGAGCTGACCTGACGACCA."

# BASE COUNT

2 a 10 c 9 g 10 t

**Query Match** 55.8%; Score 13.4; DB 10; Length 31;  
**Best Local Similarity** 73.9%; Pred. No. 2e+05;  
**Matches** 17; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 cgacaatggaataaacagctgcg 23  
 Db 25 GGGCCATGAGAAACAGCACTC 3

**RESULT** 8  
**LOCUS** A1026096 46 bp mRNA EST 27-AUG-1998  
**DEFINITION** ov94h09.s1 Soares, J., et al. Homo sapiens cDNA clone IMAGE:1645025 3' similar to TR:062006 O62006 OPA REPEAT; contains element L1 repetitive element.; mRNA sequence.  
**ACCESSION** A1026096  
**VERSION** A1026096.1 GI:3241709  
**KEYWORDS** EST.  
**SOURCE** human.  
**ORGANISM** Homo sapiens  
**REFERENCE** 1 (bases 1 to 46)  
 Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo.  
**AUTHORS** NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
**TITLE** National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index



0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD2 (g11473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."									
BASE COUNT	1 a	5 c	3 g	18 t					
ORIGIN									
Query Match	55.0%;	Score 13.2;	DB 13;	Length 27;					
Best Local Similarity	83.3%;	Pred. No. 2.4e+05;							
Matches 15;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;					
Oy	1	cgacaaatggaanaacag	18						
Db	21	CGAAAAACGGAAGAG	4						
RESULT 12									
LOCUS	C21091	38 bp	mRNA	EST	23-OCT-1996				
DEFINITION	HMGSO002607 Human adult (K.Okubo) Homo sapiens CDNA 3', mRNA sequence.								
ACCESSION	C21091								
VERSION	C21091.1	GI:1622201							
KEYWORDS	EST.								
SOURCE	human.								
ORGANISM	Homo sapiens								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.								
AUTHORS	1 (bases 1 to 38)								
TITLE	Okubo,K.								
JOURNAL	BodyMap: human gene expression database								
COMMENT	Unpublished (1995)								
	Contact: Okubo,K.								
	Institute for Molecular and Cellular Biol								
	Osaka University								
	1-3,Yamada-oka, Suita, Osaka Pref. 565, Japan								
	Tel: 06-877-5111(ex.3315)								
	Email: kousakueimcb.osaka-u.ac.jp								
	Human Gene Signature, 3'-directed CDNA sequence. We are not submitting the same CDNA sequence redundantly to DBJ since 1993.								
	For the abundance information of clones with this sequence in this library and as well as in other 3'-directed libraries, see								
	http://www.imcb.osaka-u.ac.jp/bodymap/. The sequences of the clones represented by this GS sequences is also found there.								
FEATURES	Location/Qualifiers								
source	1..38								
	/organism="Homo sapiens"								
	/db_xref="taxon:9606"								
	/clone_lib="Human adult (K.Okubo)"								
	/dev_stage="adult"								
BASE COUNT	16 a	6 c	6 g	10 t					
ORIGIN									
Query Match	55.0%;	Score 13.2;	DB 11;	Length 38;					
Best Local Similarity	83.3%;	Pred. No. 2.3e+05;							
Matches 15;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;					
Oy	5	aaatgaaaaaacgctcg	22						
Db	21	AAATGTAAAAACACCTGG	38						

RESULT 13  
AA666844 LOCUS  
DEFINITION vm65g11.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone  
IMAGE:1003172 5', mRNA sequence.  
AA666844  
VERSION AA666844.1 GI:2625545  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 59)  
REFERENCE 1 (bases 1 to 59)  
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
Waterston,R.  
TITLE The WashU-HMT Mouse EST Project  
JOURNAL Unpublished (1996)  
COMMENT Contact: Maria M/Mouse EST Project  
WashU-HMT Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@wustl.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
MGI:567388.

FEATURES  
source Location/Qualifiers  
1..59  
/organism="Mus musculus"  
/strain="B6D2 F1/J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1003172"  
/clone\_lib="Knowles Solter mouse 2 cell"  
/tissue\_type="embryo"  
/dev\_stage="2-cell"  
/lab\_host="DH10B"  
/note="Organ: embryo; Vector: pBluescribe (modified);  
Site\_1: MluI; Site\_2: SalI; Cloned unidirectionally from  
mRNA prepared from 13,500 2-cell stage embryos. Primer:  
SalI(dT): 5'-CGGTGACCGTGCACCTTTTCTTTT-3', CDNAS  
were cloned into the MluI/SalI sites of a modified  
pBluescribe vector using commercial linkers (NEB).  
Average insert size: 1.2 kb."

BASE COUNT 19 a 11 c 14 g 15 t  
ORIGIN

Query Match 55.0%; Score 13.2; DB 10; Length 59;  
Best Local Similarity 83.3%; Pred. No. 2, le+05;  
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 4 caaatggaataaacagctc 21  
|||||  
Db 2 CAAATGGAATTCACGCTC 19

RESULT 14  
AT996547 LOCUS  
DEFINITION AI996547 60 bp mRNA EST 08-SEP-1999  
CDNA clone 701667389 A. thaliana, Columbia Col-0, root-1 Arabidopsis thaliana  
AI996547  
VERSION AI996547.1 GI:5843452  
KEYWORDS EST.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 60)  
AUTHORS Chen,J., Momiyama,M., Chan,E., Mooney,M., Carroon,B., Gilliland,D.,  
Wang,X., Hillman,J., Guegler,K., Kim,C., Doyle,M., Brzoska,P.,  
Gorgone,G., Burns,D., Griffin,J., Mounanoutou,M., Nguyen,D., Tan,R.,  
Rose,M., Warren,B., Ton,B., Kasuriy,K., Borillo,C., Carpio,T.,  
Policky,J., Suzuki,G., Argentele,C., Shah,S., Nobrega,A., Murry,L.,  
Turner,C., Krikorian,S., Elder,L. and Hanson,D.  
TITLE Arabidopsis thaliana Gene Expression Microarray  
JOURNAL Unpublished (1999)  
COMMENT Contact: David Smoller, Ph.D.  
Genome Systems, Inc., a wholly owned subsidiary of Incyte  
Pharmaceuticals, Inc.  
4633 World Parkway Circle, St. Louis, MO 63134, USA  
Tel: 877-577-2733  
Fax: 314-427-3324  
Email: service@genomesystems.com.

FEATURES  
source Location/Qualifiers  
1..60  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia Col-0"  
/db\_xref="taxon:3702"  
/clone="701667389"  
/clone\_lib="A. thaliana, Columbia Col-0, root-1"  
/tissue\_type="root"  
/dev\_stage="4 - 7 weeks"  
/note="Vector: pSPORT. Site\_1: NotI; Site\_2: SalI; cDNA  
library was derived from untreated root tissue from  
Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks.  
Plants were grown in 1:1:1 peat moss/vermiculite/perlite  
soil at 22 deg. C +/- 3 deg. C under constant light, and  
watered with fertilizer. cDNA synthesis was initiated  
using a NotI-oligo(dT) primer. Double-stranded cDNA was  
blunted, ligated to SalI adaptors, digested with NotI,  
size-selected, and cloned into the NotI and SalI sites of  
the pSPORT vector."

BASE COUNT 23 a 12 c 1 g 20 t 4 others  
ORIGIN

Query Match 55.0%; Score 13.2; DB 10; Length 60;  
Best Local Similarity 78.9%; Pred. No. 2, le+05;  
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 cgacaaatggaataaacagc 19  
|||||  
Db 21 CNAACAATCCAAACATC 39

RESULT 15  
AA476000 LOCUS  
DEFINITION vh26a02.r1 Soares mammary\_gland\_NDMMG Mus musculus cDNA clone  
IMAGE:876554 5' similar to SW:CCAA MOUSE P01863 IG GAMMA-2A CHAIN C  
REGION, A ALLELE. ;, mRNA sequence.  
AA476000  
VERSION AA476000.1 GI:2203851  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 34)  
REFERENCE 1 (bases 1 to 34)  
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
Waterston,R.  
TITLE The WashU-HMT Mouse EST Project  
JOURNAL Unpublished (1996)  
COMMENT Contact: Maria M/Mouse EST Project  
WashU-HMT Mouse EST Project  
Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@watson.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:516034

Trace considered overall poor quality  
 Possible reversed clone: similarity on wrong strand  
 Seq primer: -28m13 rev2 ET from Amersham  
 High quality sequence stop: 1.

## FEATURES

source

1..34

Location/Qualifiers

/organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:876554"  
 /clone\_1lb="Soares\_mammary\_gland\_NDMG"  
 /sex="male"  
 /tissue\_type="mammary\_gland"  
 /dev\_stage="4 weeks"  
 /lab\_host="DH10B"  
 /note="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia  
 ) with a modified polylinker; Site\_1: Not I; Site\_2: Eco  
 RI; 1st strand cDNA was primed with a Not I - oligo(dT)  
 primer (5'  
 TCTTACCAATCTGAGTGGAGCGCGCGGATGGTTTTTTTTTTTTTTTTTT  
 T 3'); double-stranded cDNA was ligated to Eco RI  
 adaptors (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of the modified pT73 vector.  
 RNA provided by Dr. Minoru Ko, Wayne State Univ. Library  
 constructed and normalized by Bento Soares and M.Felima  
 Bonaldo."

BASE COUNT 7 a 12 c 5 g 10 t  
 ORIGIN

## Query Match

Best Local Similarity 54.2%; Score 13; DB 10; Length 34;  
 Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 gacaaatggaagaacagctcg 22  
 ||||| ||||| ||||| |||||  
 DB 28 GACACGTGATGACAGACATCG 8

Search completed: March 9, 2002, 00:09:10  
 Job time: 11026 sec